Short recurrences for computing extended Krylov bases for Hermitian and unitary matrices

Clara Mertens
Raf Vandebril

Report TW 641, January 2014

KU Leuven
Department of Computer Science
Celestijnenlaan 200A – B-3001 Heverlee (Belgium)

Short recurrences for computing extended Krylov bases for Hermitian and unitary matrices

Clara Mertens
Raf Vandebril

Report TW 641, January 2014

Department of Computer Science, KU Leuven

Abstract

It is well known that the projection of a matrix A onto a Krylov subspace span $\left\{\mathbf{h}, A\mathbf{h}, A^2\mathbf{h}, \ldots, A^{k-1}\mathbf{h}\right\}$ results in a matrix of Hessenberg form. We show that the projection of the same matrix A onto an extended Krylov subspace, which is of the form span $\left\{A^{-k_r}\mathbf{h}, \ldots, A^{-2}\mathbf{h}, A^{-1}\mathbf{h}, \mathbf{h}, A\mathbf{h}, A^2\mathbf{h} \ldots, A^{k_\ell}\mathbf{h}\right\}$, is a matrix of so-called extended Hessenberg form which can be characterized uniquely by its QR-factorization. In case A is a Hermitian or unitary matrix, this extended Hessenberg matrix is banded, resulting in short recurrence relations. For the unitary case, coupled two term recurrence relations are derived of which the coefficients capture all information necessary for a sparse factorization of the corresponding extended Hessenberg matrix. This generalizes the approach used by Watkins to retrieve the CMV-form for unitary matrices.

Keywords : extended Krylov, Hermitian matrices, unitary matrices, matrix functions.

Noname manuscript No.
(will be inserted by the editor)

Short recurrences for computing extended Krylov bases for Hermitian and unitary matrices

Clara Mertens · Raf Vandebril

the date of receipt and acceptance should be inserted later

Abstract It is well known that the projection of a matrix A onto a Krylov subspace span$\{\mathbf{h}, A\mathbf{h}, A^2\mathbf{h}, \ldots, A^{k-1}\mathbf{h}\}$ results in a matrix of Hessenberg form. We show that the projection of the same matrix A onto an extended Krylov subspace, which is of the form span$\left\{A^{-k_r}\mathbf{h}, \ldots, A^{-2}\mathbf{h}, A^{-1}\mathbf{h}, \mathbf{h}, A\mathbf{h}, A^2\mathbf{h} \ldots, A^{k_\ell}\mathbf{h}\right\}$, is a matrix of so-called extended Hessenberg form which can be characterized uniquely by its QR-factorization. In case A is a Hermitian or unitary matrix, this extended Hessenberg matrix is banded, resulting in short recurrence relations. For the unitary case, coupled two term recurrence relations are derived of which the coefficients capture all information necessary for a sparse factorization of the corresponding extended Hessenberg matrix. This generalizes the approach used in [15] to retrieve the CMV-form for unitary matrices.

Keywords extended Krylov · Hermitian matrices · unitary matrices · matrix functions

1 Introduction

In this article we are interested in extended Krylov subspaces, i.e., spaces spanned by successive vectors, starting from $\mathbf{h}$, out of the bilateral sequence

$$\ldots,\ A^3\mathbf{h},\ A^2\mathbf{h},\ A\mathbf{h},\ \mathbf{h},\ A^{-1}\mathbf{h},\ A^{-2}\mathbf{h},\ A^{-3}\mathbf{h},\ \ldots. \tag{1}$$

The main incentive of this article is the desire to efficiently compute orthonormal bases (of which the vectors are stored in a unitary matrix Q) of such extended Krylov subspaces. The matrix H satisfying $AQ = QH$ is a structured matrix, which will be refered to as a matrix of extended Hessenberg form, and which we also aim to compute in an efficient way.

Extended Krylov subspaces are fundamental tools in numerous applications including matrix functions [9], model order reduction [1,4] and Lyapunov equations [10], among others. For example,

The research was partially supported by the Research Council KU Leuven, projects OT/11/055 (Spectral Properties of Perturbed Normal Matrices and their Applications), CREA/13/012 (Can Unconventional Eigenvalue Algorithms Supersede the State-Of-The-Art), PFV/10/002 (Optimization in Engineering, OPTEC) , by the Fund for Scientific Research–Flanders (Belgium) project G034212N (Reestablishing Smoothness for Matrix Manifold Optimization via Resolution of Singularities), and by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office, Belgian Network DYSCO (Dynamical Systems, Control, and Optimization).

Clara Mertens
KU Leuven, Dept. of Computer Science, 3000 Leuven, Belgium
E-mail: clara.mertens@cs.kuleuven.be

Raf Vandebril
KU Leuven, Dept. of Computer Science, 3000 Leuven, Belgium
E-mail: raf.vandebril@cs.kuleuven.be

in [8] short recurrences for extended Krylov subspaces are investigated, relating them to Laurent polynomials, in order to evaluate expressions of the form $f(A)\mathbf{v}$ with A a large symmetric matrix and $\mathbf{v}$ a column vector.

The article is organized as follows. In Section 2 extended Hessenberg matrices are characterized by means of their QR-factorization, the matrix Q exhibiting a zigzag shaped pattern of rotations [2,12,14]. In Section 3 it is established that the matrix H satisfying $AQ = QH$, with Q a unitary matrix, is a matrix of extended Hessenberg form. The remainder of the paper is focused on Hermitian and unitary matrices as they exhibit a special structure giving rise to short recurrences. In Section 3 we show that the matrix framework built in Section 2 easily allows us to predict the banded structure of Hermitian and unitary extended Hessenberg matrices. This generalizes part of the theoretical results established in [8] for Hermitian matrices. Section 4 comprises a new algorithm for computing an orthonormal basis for an arbitrary extended Krylov subspace of a unitary matrix. Given the order in which the vectors are chosen from the bilateral sequence (1), the algorithm returns an orthonormal basis for the corresponding extended Krylov subspace together with a sparse factorization of the associated extended Hessenberg matrix H. The latter captures the coefficients of the recurrence relations between the orthonormal basis vectors. The orthonormal basis is retrieved recursively making use of coupled two term recurrences. The same approach of coupled two term recurrences was used in [15] to obtain the so-called CMV-shape, the latter being a specific extended Hessenberg matrix which corresponds to the extended Krylov space consisting of an alternating sequence of positive and negative powers of the matrix A. However, this article provides a generalization of this approach which is applicable to any extended Krylov subspace. In Section 5 longer recurrences are considered, of which the length depends on the number of successive positive or negative powers of the matrix A.

The following notation is used throughout the article. Matrices are denoted by upper case letters $A = (a_{ij})$, the element in the matrix A on the intersection of the ith row and jth column is given by a_{ij}, and with I the identity matrix is signaled. Vectors are written in bold face and lower case letters, e.g., $\mathbf{x}, \mathbf{y}$, and $\mathbf{z}$. With $\operatorname{span}\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ the subspace generated by vectors $\mathbf{x}, \mathbf{y}, \mathbf{z}$ is meant. The standard inner product is denoted as $\langle \mathbf{x}, \mathbf{y}\rangle = \mathbf{y}^*\mathbf{x}$, with $\cdot^*$ the Hermitian conjugate. Throughout the article, Matlab's indexing is used to establish the location of submatrices.

2 Properties of extended Hessenberg matrices

It is well known that the projection of a matrix onto a Krylov subspace is a matrix of Hessenberg form. It will be shown in Section 3 that this statement can be generalized to extended Krylov subspaces, i.e., the projection of a matrix onto an extended Krylov subspace will be a matrix of extended Hessenberg form. Therefore, extended Hessenberg matrices will play a key role in this article. The matrices under consideration look like the ones in Figure 2, possessing diagonal blocks of either Hessenberg or inverse Hessenberg form.

Definition 1 Given an ordered list of indices $i_1 = 1 < i_2 < i_3 < \ldots < i_{m-1} < i_m = n$ with n the dimension of the matrix H. Then H is called an extended Hessenberg matrix if it is of block upper triangular form where the blocks $H(i_j : i_{j+1}+1, i_j : i_{j+1}+1)$ form a sequence of matrices alternatingly of Hessenberg and inverse Hessenberg form, each block sharing a 2×2 submatrix with the next block.

Without loss of generality and for the ease of explanation we assume in the remainder of the text the upper left diagonal block to be of Hessenberg form. Then we have that the blocks $H(i_j : i_{j+1}+1, i_j : i_{j+1}+1)$, for odd j are of Hessenberg form and for even j of inverse Hessenberg form.

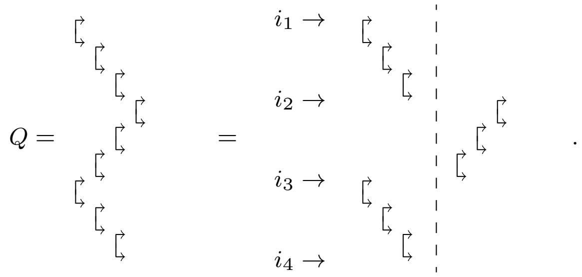
These matrices were termed *compressed matrices* in [14], as they admit an equally expensive QR-factorization as a Hessenberg and inverse Hessenberg matrix. The QR-factorizations of these matrices are comprised of $n-1$ rotations and an upper triangular matrix, the positioning of the individual rotations characterizing the matrix. Figure 1 depicts the *detailed factorizations* of a Hessenberg matrix and a matrix of inverse Hessenberg form[1]. Each bracket stands for a rotation, with arrowheads targeting the rows affected by the rotation.

(a) matrix of Hessenberg form

(b) matrix of inverse Hessenberg form

Fig. 1: Graphical depiction of two detaileed QR-factorizations.

A Hessenberg matrix has a *descending* sequence of rotations, whereas a matrix of Hessenberg form exhibits an *ascending* sequence of rotations. However, for the detailed factorizations of a generic extended matrix H no specific ordering is imposed. An efficient QR-factorization of such matrix can easily be computed and equals $Q_{p_1} \dots Q_{p_{n-1}} R$ with $p_1, \dots, p_{n-1}$ a permutation of $1, \dots, n-1$. The ensemble of rotations exhibits thus sort of twisted shape and can be decomposed as a product VW^H with both V and W of unitary Hessenberg form. Roughly one can state that the matrix V encompasses the rotations associated to the Hessenberg blocks in H (descending ordering) and W the rotations linked to inverse Hessenberg blocks (ascending ordering). This *double Hessenberg* factorization was presented in a more general context in [12] and (2) displays such a factorization for a particular twisted shape. The indices i_j as defined in the first paragraph are also displayed; in constrast to the full matrix, these indices point here precisely to the beginning of each individual Hessenberg or inverse Hessenberg block. On the left of the dashed line a unitary Hessenberg matrix composed of diagonal unitary Hessenberg blocks is observed, on the right a unitary inverse Hessenberg with non compatible block diagonal structure is shown.



(2)

This factorization presents another viable manner of detailing the structure of H. The matrix H has diagonal blocks $H(i_j : i_{j+1}+1, i_j : i_{j+1}+1)$ (for odd j) of Hessenberg form; and the submatrices $H(i_j : i_{j+1}, i_j : i_{j+1})$ (for even j) of inverse Hessenberg form. Every inverse Hessenberg matrix is followed by a row $H(i_{j+1}+1, :)$ which is a multiple of the last row $H(i_{j+1}, :)$ of this inverse Hessenberg matrix. This extra row is induced by the left multiplication with a Hessenberg matrix. Figure 2 portrays both interpretations.

Lemma 1 *The inverse of an extended Hessenberg matrix is again an extended Hessenberg matrix with the Hessenberg and inverse Hessenberg blocks swapped.*

Inverting the double Hessenberg factorization and passing the rotations from right to left through the upper triangular matrix[2] proves the statement. Figure 3 illustrates the block structure of an

[1] Throughout the article, we will refer to an inverse Hessenberg matrix or a matrix of inverse Hessenberg form, as a matrix exhibiting a QR-factorization like the one in Figure 1(b). In case the matrix is nonsingular, it is the inverse of a Hessenberg matrix, justifying this choice of terminology.

[2] Given a rotation Q_{p_i}, the product $Q_{p_i} R$ can be rewritten as $\tilde{R}\tilde{Q}_{p_i}$, where $\tilde{R}$ is upper triangular. Hence the rotation has been passed from the left to the right of the upper-triangular matrix.

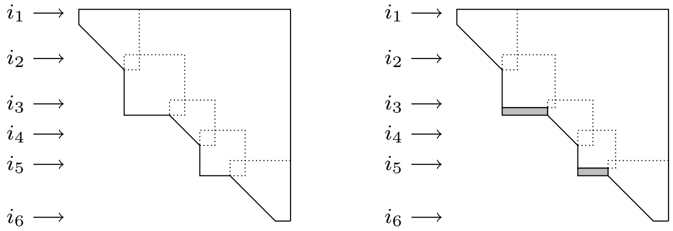


Fig. 2: Two interpretations of the extended matrix structure.

extended Hessenberg matrix, shown in the leftmost graphic, and the two interpretations of its inverse, shown in the middle and rightmost graphic of Figure 3. Special attention is required when examining the overlapping elements.

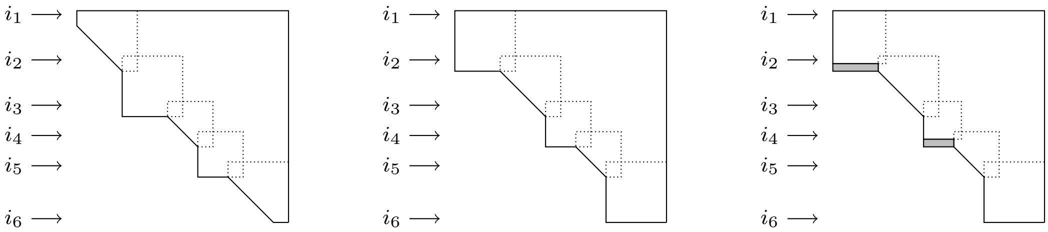


Fig. 3: Structure of an extended matrix and its inverse.

3 Extended Krylov spaces

An extended Krylov subspace is generated by not only powers of A but also of the inverse A^{-1}. By construction, a k dimensional extended Krylov subspace is spanned by k successive vectors, starting with $\mathbf{h}$, out of the bilateral sequence

$$\ldots,\ A^3\mathbf{h},\ A^2\mathbf{h},\ A\mathbf{h},\ \mathbf{h},\ A^{-1}\mathbf{h},\ A^{-2}\mathbf{h}, A^{-3}\mathbf{h},\ \ldots. \tag{3}$$

The order in which the vectors are added to the subspace is crucial, and recorded in a *position* vector $\mathbf{p}$ of length $n-2$, n the dimension of A, comprised of characters ℓ and r. The ℓ indicates that the next vector in the subspace is taken on the left, the r points out that the next vector is taken on the right of (3), e.g., the CMV-shape corresponds to the position vector $\mathbf{p} = [\ell, r, \ell, r, \ell, r, \ldots]$. By positioning positive powers of A in the bilateral sequence on the left, and inverse powers on the right we are consistent with [12,14]. The nth vector spans the extended Krylov subspace together with the vectors previously selected out of the bilateral sequence (3), and is thus independent of the position vector $\mathbf{p}$, which explains the fact that the vector $\mathbf{p}$ is of length $n-2$, rather than $n-1$.

Suppose that in the first $k-1$ components of $\mathbf{p}$ the symbol ℓ appears k_ℓ times and r appears k_r times ($k_r + k_\ell = k-1$), then

$$\mathcal{K}_{\mathbf{p},k}(A,\mathbf{h}) = \operatorname{span}\Big\{A^{-k_r}\mathbf{h}, \ldots, A^{-2}\mathbf{h}, A^{-1}\mathbf{h}, \mathbf{h}, A\mathbf{h}, A^2\mathbf{h}, \ldots, A^{k_\ell}\mathbf{h}\Big\}, \tag{4}$$

depicts the extended Krylov space of dimension $k = k_r + k_\ell + 1$. Clearly

$$\mathcal{K}_{\mathbf{p},k}(A,\mathbf{h}) = \operatorname{span}\Big\{A^{-k_r}\mathbf{h}, \ldots, A^{k_\ell}\mathbf{h}\Big\} = A^{-k_r}\mathcal{K}_k(A,\mathbf{h}) = A^{k_\ell}\mathcal{K}_k(A^{-1},\mathbf{h}).$$

The vectors $\mathbf{q}_1, \mathbf{q}_2, \ldots \mathbf{q}_n$ are constructed by successively orthonormalizing the columns of the Krylov matrix. The following table presents in a comprehensible way the essential relations. Assume for now $p_1 = \ell$. The first and second row shows the position vector, the third row the vector to be orthonormalized, and the trailing row depicts the computed orthogonal vectors.

$\mathbf{p}$		p_1	p_2		p_{i_2-1}	p_{i_2}		p_{i_3-1}	p_{i_3}	
p_i		ℓ	ℓ	$\ldots$	ℓ	r	$\ldots$	r	ℓ	$\ldots$
columns of $\mathcal{K}$	$\mathbf{h}$	$A\mathbf{h}$	$A^2\mathbf{h}$		$A^{i_2-1}\mathbf{h}$	$A^{-1}\mathbf{h}$		$A^{i_2-i_3}\mathbf{h}$	$A^{i_2}\mathbf{h}$	
columns of Q	$\mathbf{q}_1$	$\mathbf{q}_2$	$\mathbf{q}_3$		$\mathbf{q}_{i_2}$	$\mathbf{q}_{i_2+1}$		$\mathbf{q}_{i_3}$	$\mathbf{q}_{i_3+1}$	

$\mathbf{p}$		p_{i_4-1}	p_{i_4}		p_{i_5-1}	p_{i_5}	
p_i	$\ldots$	ℓ	r	$\ldots$	r	ℓ	$\ldots$
columns of $\mathcal{K}$		$A^{i_2-i_3+i_4-1}\mathbf{h}$	$A^{i_2-i_3-1}\mathbf{h}$		$A^{i_2-i_3-i_4+i_5+1}\mathbf{h}$	$A^{i_2-i_3+i_4}\mathbf{h}$	
columns of Q		$\mathbf{q}_{i_4}$	$\mathbf{q}_{i_4+1}$		$\mathbf{q}_{i_5}$	$\mathbf{q}_{i_5+1}$	

A non-interrupted sequence of vectors out of the two trailing rows starting with $\mathbf{h}$ and $\mathbf{q}_1$ generate, by construction, identical subspaces:

$$\text{span}\Big\{A^{-k_r}\mathbf{h}, \ldots, A^{k_\ell}\mathbf{h}\Big\} = \text{span}\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_k\}.$$

The inclusion $A\mathcal{K}_k(A, \mathbf{h}) \subset \mathcal{K}_{k+1}(A, \mathbf{h})$ for standard Krylov subspaces is a key item in the derivation of the Hessenberg structure capturing the recurrence coefficients for the orthogonal vectors. For the extended case, the values of the position vector affect the relations. This is captured in Lemma 2.

Lemma 2 (Lemma 3.6 in [14]) *For $k = 1, \ldots, n-2$,*

(a) if $p_k = \ell$, then $A\mathcal{K}_{\mathbf{p},k}(A, \mathbf{h}) \subseteq \mathcal{K}_{\mathbf{p},k+1}(A, \mathbf{h})$.
(b) if $p_k = r$, then $A^{-1}\mathcal{K}_{\mathbf{p},k}(A, \mathbf{h}) \subseteq \mathcal{K}_{\mathbf{p},k+1}(A, \mathbf{h})$.

3.1 Recursions for the arbitrary matrix case

Once the position vector is set, we are interested in the matrix H satisfying $AQ = QH$, Q unitary. In [12, 14] the structure of the matrix H was already established for the generic matrix case. In the spirit of this article, we will, however, deduce an alternative proof along the same lines as for the Hessenberg setting. The matrix H will be of extended Hessenberg form, which is of block diagonal form with the blocks of Hessenberg and inverse Hessenberg form sharing some of the upper left and lower right elements. More precisely, based on a given position vector $\mathbf{p}$ we define the characterizing indices (see Section 2) i_j as follows: $i_1 = 1$ and i_j is the position k such that the $(j-1)$th transition from ℓ to r, or r to ℓ takes place between p_{k-1} and p_k. The list is closed by a trailing $i_j = n$, where n is the dimension of A. We will prove now that the matrix H meets the extended Hessenberg matrix constraints. To do so, we rely on the following result:

Lemma 3 (Corollary of the Nullity Theorem [13]) *Suppose $A \in \mathbb{R}^{n\times n}$ is a nonsingular matrix, and α, β are nonempty subsets of N with $|\alpha| < n$ and $|\beta| < n$, N being the index set $\{1, 2, \ldots, n\}$. Then*

$$\mathit{rank}\left(A^{-1}(\alpha; \beta)\right) = \mathit{rank}\left(A(N\backslash\beta; N\backslash\alpha)\right) + |\alpha| + |\beta| - n.$$

Reconsidering the relation $AQ = QH$ of which we want to extract the structure of H, we see that for $i_1 = 1 \leq i \leq i_2 - 1$, $A\mathbf{q}_i$ is a linear combination of the vectors $\mathbf{q}_1, \ldots, \mathbf{q}_{i+1}$. This statement results also from Lemma 2: the values of $p_1, \ldots, p_{i_2-1}$ all equal ℓ and therefore for $1 \leq i \leq i_2 - 1$

$$A[\mathbf{q}_1, \ldots, \mathbf{q}_i] = [\mathbf{q}_1, \ldots \mathbf{q}_{i+1}]H(1 : i+1, 1 : i),$$

with $H(1:i_2+1,1:i_2+1)$ thus of Hessenberg form. The conclusion on the structure can be extended to all orthogonal vectors $\mathbf{q}_i$ for which $p_i = \ell$. More precisely, Lemma 2 is also valid for all vectors $\mathbf{q}_{i_j},\ldots,\mathbf{q}_{i_{j+1}-1}$, with odd j, implying the general relations

$$A[\mathbf{q}_1,\ldots,\mathbf{q}_i] = [\mathbf{q}_1,\ldots,\mathbf{q}_{i+1}]H(1:i+1,1:i),$$

where $i_j \leq i \leq i_{j+1}-1$, j odd, imposing thus a Hessenberg structure on the blocks $H(i_j : i_{j+1}+1, i_j : i_{j+1}+1)$ for odd j.

Unfortunately there is not much to say about the other diagonal blocks. Lemma 2 only implies that for $i_j \leq i \leq i_{j+1}$, j even, with $p_{i_{j+1}}$ thus the earliest next appearance of ℓ that

$$A\mathcal{K}_{\mathbf{p},i} \subseteq \mathcal{K}_{\mathbf{p},i_{j+1}}.$$

This yields

$$A[\mathbf{q}_1,\ldots,\mathbf{q}_i] = [\mathbf{q}_1,\ldots,\mathbf{q}_{i_{j+1}+1}]H(1:i_{j+1}+1,1:i),$$

and we get thus a dense block $H(i_j : i_{j+1}+1, i_j : i_{j+1}+1)$ for even j.

To prove the inverse Hessenberg structure of the dense diagonal blocks we use the inverse relation $A^{-1}Q = QH^{-1}$. First we prove that the blocks $H^{-1}(i_j : i_{j+1}, i_j : i_{j+1})$ are of Hessenberg form, for even. To construct the matrix Q starting with A^{-1} the elements in the position vector need to be exchanged, i.e., symbols r become ℓ's and vice versa. In this way exactly the same extended Krylov matrix is built leading thus to an identical matrix Q. Identical reasoning as in the previous paragraph leads us to the Hessenberg structure of the diagonal blocks under consideration. Once the Hessenberg structure of the blocks $H^{-1}(i_j : i_{j+1}+1, i_j : i_{j+1}+1)$ is established, Lemma 3 implies that

$$\begin{aligned}\operatorname{rank}\left(H(i_j+k:n,1:i_j+l)\right) &= \operatorname{rank}\left(H^{-1}(i_j+l+1:n,1:i_j+k-1)\right)+l-k+1,\\ &= (k-l)+(l-k+1) = 1,\end{aligned}$$

with $0 \leq l \leq k \leq i_{j+1}-i_j$, the last equality being due to the fact that $H^{-1}(i_j : i_{j+1}+1, i_j : i_{j+1}+1)$ is a Hessenberg matrix. As

$$\operatorname{rank}\left(H(i_j+k:n,1:i_j+l)\right) = \operatorname{rank}\left(H(i_j+k:i_{j+1},i_j:i_j+l)\right),$$

it follows that all submatrices of $H(i_j : i_{j+1}+1, i_j : i_{j+1}+1)$, j even, below the first superdiagonal and including the bottom left entry are of rank one, establishing the inverse Hessenberg structure.

3.2 Hermitian matrices

For Hermitian matrices A the matrix H satisfying $AQ = QH$, is banded. The bandwidth depends on the structure of the position vector $\mathbf{p}$. The larger the difference $i_{j+1}-i_j$, the larger the bandwidth. In [8] this result was established by making use of orthogonal Laurent polynomials. However, combining the result of Subsection 3.1, which predicts the structure of the matrix H as a function of the position vector $\mathbf{p}$, and the relation $H = H^*$, the banded structure of the matrix H follows at once, without the need of any computation.

Depending on the bandwidth, recurrences of limited length are obtained. Figure 4 shows the structure of a general extended Hessenberg matrix together with its inverse. Projecting the Hermitian conjugate of both matrices on themselves, the zero structure as depicted in Figure 5 is found. It is assumed that the position vector $\mathbf{p}$ starts with the symbol ℓ.

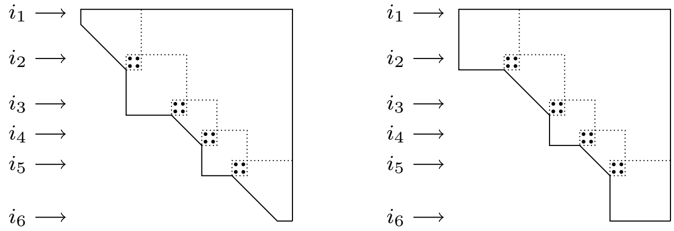


Fig. 4: Structure of a general extended Hessenberg matrix H satisfying $AQ = QH$ and its inverse.

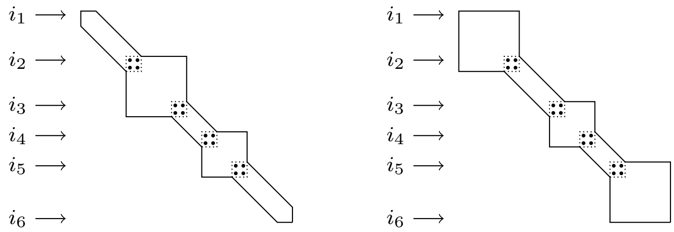


Fig. 5: Structure of an Hermitian extended Hessenberg matrix H satisfying $AQ = QH$ and its inverse.

The values of $p_{i_j}, p_{i_j+1}, \cdots, p_{i_{j+1}-1}$, for j odd, all equal ℓ. Hence Lemma 2(a) and the relation $H^* = H$ imply the recurrence relations

$$\beta \mathbf{q}_{i+1} = A\mathbf{q}_i - \sum_{k=i-1}^{i} \alpha_k \mathbf{q}_k, \qquad i_n + 2 \leq i \leq i_{n+1} - 1, \tag{5}$$

with $\alpha_k = \langle A\mathbf{q}_k, \mathbf{q}_k \rangle$, $\beta = \sqrt{||A\mathbf{q}_i||_2 - \sum_{k=i-1}^{i} |\alpha_k|^2}$,

$$\gamma \mathbf{q}_{i_j+1} = A\mathbf{q}_{i_j} - \sum_{k=i_{j-1}}^{i_n} \delta_k \mathbf{q}_k, \tag{6}$$

with $\delta_k = \langle A\mathbf{q}_{i_j}, \mathbf{q}_k \rangle$, $\gamma = \sqrt{||A\mathbf{q}_{i_j}||_2 - \sum_{k=i_{j-1}}^{i_j} |\delta_k|^2}$, and

$$\varepsilon \mathbf{q}_{i_j+2} = A\mathbf{q}_{i_j+1} - \sum_{k=i_{j-1}}^{i_j+1} \chi_k \mathbf{q}_k, \tag{7}$$

with $\chi_k = \langle A\mathbf{q}_{i_j+1}, \mathbf{q}_k \rangle$, $\varepsilon = \sqrt{||A\mathbf{q}_{i_j+1}||_2 - \sum_{k=i_{j-1}}^{i_j+1} |\chi_k|^2}$.

The values of $p_{i_j}, p_{i_j+1}, \cdots, p_{i_{j+1}-1}$, for j even, all equal r. Hence, by making use of Lemma 2(b) and $\left(H^{-1}\right)^* = H^{-1}$ the same formulas as when n is odd apply with A replaced by A^{-1}. We obtain

$$\tilde{\beta} \mathbf{q}_{i+1} = A^{-1}\mathbf{q}_i - \sum_{k=i-1}^{i} \tilde{\alpha}_k \mathbf{q}_k, \qquad i_n + 2 \leq i \leq i_{n+1} - 1, \tag{8}$$

with $\tilde{\alpha}_k = \langle A^{-1}\mathbf{q}_i, \mathbf{q}_k\rangle$, $\tilde{\beta} = \sqrt{||A^{-1}\mathbf{q}_i||_2 - \sum_{k=i-1}^{i} |\tilde{\alpha}_k|^2}$,

$$\tilde{\gamma}\mathbf{q}_{i_j+1} = A^{-1}\mathbf{q}_{i_j} - \sum_{k=i_{j-1}}^{i_j} \tilde{\delta}_k \mathbf{q}_k, \tag{9}$$

with $\tilde{\delta_k} = \langle A^{-1}\mathbf{q}_{i_j}, \mathbf{q}_k\rangle$, $\tilde{\gamma} = \sqrt{||A^{-1}\mathbf{q}_{i_j}||_2 - \sum_{k=i_{j-1}}^{i_j} |\tilde{\delta}_k|^2}$, and

$$\tilde{\varepsilon}\mathbf{q}_{i_j+2} = A^{-1}\mathbf{q}_{i_j+1} - \sum_{k=i_{j-1}}^{i_j+1} \tilde{\chi}_k \mathbf{q}_k, \tag{10}$$

with $\tilde{\chi_k} = \langle A^{-1}\mathbf{q}_{i_j+1}, \mathbf{q}_k\rangle$, $\tilde{\varepsilon} = \sqrt{||A^{-1}\mathbf{q}_{i_j+1}||_2 - \sum_{k=i_{j-1}}^{i_j+1} |\tilde{\chi}_k|^2}$.

If the value of the first entry of the position vector $\mathbf{p}$ is r, the same recurrence relations apply with A and A^{-1} swapped.

We will now apply the above results to the examples shown in [8]. In the remainder of this section we switch to the index notation used in [8], i.e., the index m denotes that the corresponding vector is the result of orthonormalizing $A^m\mathbf{h}$ with $\mathbf{h}$ the starting vector against the previously computed orthonormal vectors. For the first example consider the position vector $\mathbf{p} = [\ell, r, \ell, r, \ell, \ldots]$. We aim to compute the corresponding orthonormal basis $\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_{-1}, \mathbf{v}_2, \mathbf{v}_{-2}, \ldots$. As in this case the indices $i_1, i_2, \ldots, i_n$ are consecutive numbers, equations (5), (7), (8) and (10) do not apply. Hence, the resulting basis is computed by alternated use of equations (6) and (9) which, using the notation of [8], transforms into the recurrence relations

$$\delta_{m+1}\mathbf{v}_{m+1} = (A - \alpha_{-m}I_n)\mathbf{v}_{-m} - \alpha_m \mathbf{v}_m, \qquad \alpha_j = \langle A\mathbf{v}_{-m}, \mathbf{v}_j\rangle,$$

$$\delta_{-m}\mathbf{v}_{-m} = (A^{-1} - \beta_m I_n)\mathbf{v}_m - \beta_{-m+1}\mathbf{v}_{-m+1}, \qquad \beta_j = \langle A^{-1}\mathbf{v}_m, \mathbf{v}_j\rangle,$$

$\delta_{m+1}, \delta_{-m}$ orthonormalization factors to make the corresponding vector of unit length.

Another example which is investigated in [8] is the position vector $\mathbf{p} = [\ell, \ell, r, \ell, \ell, r, \ldots]$, giving rise to an orthonormal basis $\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_{-1}, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_{-2}, \ldots$. Examining the position vector $\mathbf{p}$, one concludes that the corresponding basis can be computed by alternated use of equations (6), (7) and (9). Again, using the notation of [8], these equations transform into the recurrence relations

$$\delta_{2m+1}\mathbf{v}_{2m+1} = (A - \alpha_{-m,-m}I_n)\mathbf{v}_{-m} - \alpha_{-m,2m}\mathbf{v}_{2m},$$

$$\delta_{2m}\mathbf{v}_{2m} = (A - \alpha_{2m-1,2m-1}I_n)\mathbf{v}_{2m-1} - \alpha_{2m-1,-m+1}\mathbf{v}_{-m+1} - \alpha_{2m-1,2m-2}\mathbf{v}_{2m-2},$$

$$\delta_{-m}\mathbf{v}_{-m} = (A^{-1} - \beta_{2m,2m}I_n)\mathbf{v}_{2m} - \beta_{2m,-m+1}\mathbf{v}_{-m+1} - \beta_{2m,2m-1}\mathbf{v}_{2m-1},$$

with $\alpha_{j,k} = \langle A\mathbf{v}_k, \mathbf{v}_j\rangle$, $\beta_{j,k} = \langle A^{-1}\mathbf{v}_k, \mathbf{v}_j\rangle$ and δ_j a positive factor making the corresponding vector of unit length.

3.3 Unitary matrices

The class of unitary matrices in this extended Krylov space setting exhibits the most intriguing properties, as recently a lot of attention is paid to the rediscovery of the so-called CMV-shape, [5, 11,15]. The CMV-shape coincides with the pentadiagonal unitary matrix associated to an alternating appearance of ℓ and r in the position vector. Generalizations for other types of position vectors can, e.g., be found in [2].

The alternating appearance of ℓ's and r's implies $i_j = j$. The leftmost graphic in Figure 6 shows the resulting pentadiagonal structure capturing the recurrence coefficients. The nonzero matrix elements are represented by bullets. We will prove now in a different way that this structure is correct.

The second and third figure reveal the structure of the lower triangular part, in correspondence to Subsection 3.1. The elements surrounded by a dotted line belong to a Hessenberg block, the ones surrounded by a dense line make up an inverse Hessenberg matrix. The second figure shows the first interpretation of the alternating Hessenberg and inverse Hessenberg blocks, the third figure the second interpretation, the extra row is plotted on a gray shade. The inverse is characterized by a simple interchange of inverse and Hessenberg blocks. The last figure provides the structure, with marked blocks, of the inverse.

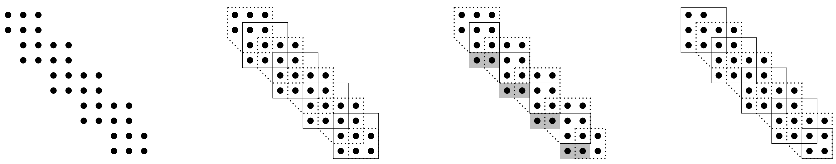


Fig. 6: CMV structure.

Projecting the structure of the inverse matrix on the upper Hessenberg part, using the relation $U^* = U^{-1}$, nicely provides us the zero structure of the upper triangular part of the matrix H.

The same reasoning applies to any type of position vector. As an example, consider the position vector $\mathbf{p} = [\ell, \ell, r, \ell, r, r, \ell, r, \ldots]$. The latter gives rise to the extended Hessenberg matrix as depicted in the leftmost graphic in Figure 7. The rightmost graphic of Figure 7 represents the inverse. Again the structure of the lower triangular part is retrieved in correspondence to Subsection 3.1, whereas the zero structure of the upper triangular part results from the relation $U^* = U^{-1}$.

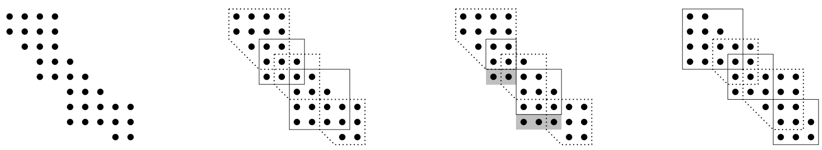


Fig. 7: Interpretations of a unitary extended Hessenberg matrix and its inverse.

4 Two term recursions for unitary matrices

In this section an algorithm will be proposed for computing an orthonormal basis for an extended Krylov subspace, based on coupled two term recurrences, i.e., vectors are computed successively by only making use of two of the previously computed vectors. In [15] it is shown that the vectors which consitute a basis for the extended Krylov subspace corresponding to the position vector $\mathbf{p} = [\ell, r, \ell, r, \ell, \ldots]$ can be computed recursively with a coupled two term recurrence. The term 'coupled' refers to the fact that the latter basis is computed simultaneously with a basis for the extended Krylov subspace corresponding to the position vector $\mathbf{p} = [r, \ell, r, \ell, \ldots]$. Projecting the unitary matrix under consideration on this extended Krylov subspace, gives rise to a matrix of CMV-form. Our method places the approach used in [15] in a more general framework, resulting in an algorithm which is able to compute an extended Krylov basis for any position vector $\mathbf{p}$.

In the first subsection we revisit the CMV-shape and apply our method to derive the coupled two term recurrences as stated in [15]. The second subsection discusses the intended generalization, illustrating its applicability in an example in the third subsection. Finally, in the fourth subsection a more general example and the corresponding algorithm are provided.

4.1 CMV shape

Given a unitary matrix U and a random starting vector $\mathbf{h}$, the sequence

$$\mathbf{h},\ U\mathbf{h},\ U^{-1}\mathbf{h},\ U^{2}\mathbf{h},\ U^{-2}\mathbf{h},\ U^{3}\mathbf{h},\ U^{-3}\mathbf{h},\ \ldots$$

corresponding to an alternating appearance of ℓ's and r's, can be orthonormalized using two term recursions. To this end, we orthonormalize this sequence simultaneously with the sequence

$$\mathbf{h},\ U^{-1}\mathbf{h},\ U\mathbf{h},\ U^{-2}\mathbf{h},\ U^{2}\mathbf{h},\ U^{-3}\mathbf{h},\ U^{3}\mathbf{h},\ \ldots .$$

We denote the orthonormal basis associated with the first sequence by $(\mathbf{v}_k)$, and that associated with the second sequence by $(\mathbf{w}_k)$. This means we will successively compute the basis vectors

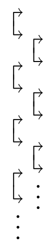
$$\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_{-1}, \mathbf{v}_2, \mathbf{v}_{-2}, \mathbf{v}_3, \mathbf{v}_{-3}, \ldots,$$

as well as

$$\mathbf{w}_0, \mathbf{w}_{-1}, \mathbf{w}_1, \mathbf{w}_{-2}, \mathbf{w}_2, \mathbf{w}_{-3}, \mathbf{w}_3, \ldots,$$

where $\mathbf{v}_k$, $\mathbf{w}_k$ denote the result of orthonormalizing $U^k\mathbf{h}$ against all previously computed vectors of $(\mathbf{v}_k)$, $(\mathbf{w}_k)$ respectively. We will now present our method as a specific instance applied to CMV, deriving the same coupled two term recurrences as in [15].

The position vector $\mathbf{p} = [\ell, r, \ell, r, \ell, r, \ldots]$ corresponds to the sequence of rotations (11).



(11)

Consider the scheme depicted in Figure 8, which consists of three sequences of vectors. After orthormalization these give rise to the bases $(\mathbf{v}_k)$, $(\tilde{\mathbf{v}}_k)$ and $(\mathbf{w}_k)$. The sequences are constructed as follows. The sequence corresponding to $(\tilde{\mathbf{v}}_k)$ is obtained by multiplying the sequence corresponding to $(\mathbf{v}_k)$ with the unitary factor U^{-1}. This immediately implies the relation $\mathbf{v}_k = U\tilde{\mathbf{v}}_{k-1}$ for all $k \in \mathbb{Z}$. Next consider the twisted pattern of rotations as depicted in (11). Intuitively, the locations of the rotations in the first column of this pattern link the sequences corresponding to the bases $(\tilde{\mathbf{v}}_k)$ and $(\mathbf{w}_k)$, and the locations of the rotations in the second column link the sequences corresponding to the bases $(\mathbf{w}_k)$ and $(\mathbf{v}_k)$. More precisely, the sequence corresponding to $(\mathbf{w}_k)$ can be derived from that of $(\tilde{\mathbf{v}}_k)$ by swapping neighbouring elements whenever there is a rotation acting on the same positions in the first column. Analogously, the sequence corresponding to $(\mathbf{v}_k)$ can be derived from that of $(\mathbf{w}_k)$ be swapping neighbouring elements whenever there is a rotation acting on the same positions in the second column. This is depicted by the arrows in Figure 8.

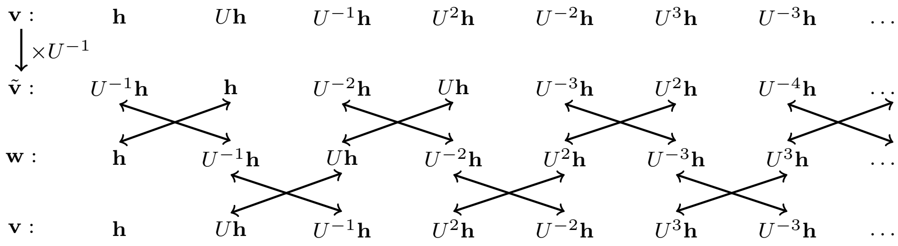


Fig. 8: Links between the sequences corresponding to the bases $(\mathbf{v}_k)$ and $(\mathbf{w}_k)$.

Note that the arrows in Figure 8 follow the same pattern as the rotations in (11) but mirrored.

Lemma 4 *Using the scheme of Figure 8, the following equalities can be derived :*

$$\mathit{span}\{\tilde{\mathbf{v}}_{-1}, \ldots, \tilde{\mathbf{v}}_{-k}, \tilde{\mathbf{v}}_{k-1}\} = \mathit{span}\{\mathbf{w}_0, \ldots, \mathbf{w}_{k-1}, \mathbf{w}_{-k}\}, \tag{12}$$
$$\mathit{span}\{\tilde{\mathbf{v}}_{-1}, \ldots, \tilde{\mathbf{v}}_{-k}, \tilde{\mathbf{v}}_{k-1}\} = \mathit{span}\{\tilde{\mathbf{v}}_{-1}, \ldots, \tilde{\mathbf{v}}_{-k}, \mathbf{w}_{k-1}\}, \tag{13}$$
$$\mathit{span}\{\mathbf{w}_0, \ldots, \mathbf{w}_{k-1}, \mathbf{w}_{-k}\} = \mathit{span}\{\mathbf{w}_0, \ldots, \mathbf{w}_{k-1}, \tilde{\mathbf{v}}_{-k}\}, \tag{14}$$

and

$$\mathit{span}\{\mathbf{w}_0, \ldots, \mathbf{w}_{-k}, \mathbf{w}_k\} = \mathit{span}\{\mathbf{v}_0, \ldots, \mathbf{v}_k, \mathbf{v}_{-k}\}, \tag{15}$$
$$\mathit{span}\{\mathbf{w}_0, \ldots, \mathbf{w}_{-k}, \mathbf{w}_k\} = \mathit{span}\{\mathbf{w}_0, \ldots, \mathbf{w}_{-k}, \mathbf{v}_k\}, \tag{16}$$
$$\mathit{span}\{\mathbf{v}_0, \ldots, \mathbf{v}_k, \mathbf{v}_{-k}\} = \mathit{span}\{\mathbf{v}_0, \ldots, \mathbf{v}_k, \mathbf{w}_{-k}\}, \tag{17}$$

for all $k > 0$.

Proof We provide a proof only for equalities (12)-(14) as equalities (15)-(17) follow analogously. Equality (12) is derived from the fact that the sequences in Figure 8 underlying the bases $(\tilde{\mathbf{v}}_k)$ and $(\mathbf{w}_k)$ are equal to each other up to permutation, and therefore span the same subspaces. More precisely, we have that

$$\begin{aligned}\operatorname{span}\{\tilde{\mathbf{v}}_{-1}, \tilde{\mathbf{v}}_0, \ldots, \tilde{\mathbf{v}}_{-k}, \tilde{\mathbf{v}}_{k-1}\} &= \operatorname{span}\{U^{-1}\mathbf{h}, \mathbf{h}, \ldots, U^{-k}\mathbf{h}, U^{k-1}\mathbf{h}\}\\ &= \operatorname{span}\{\mathbf{h}, U^{-1}\mathbf{h}, \ldots, U^{k-1}\mathbf{h}, U^{-k}\mathbf{h}\}\\ &= \operatorname{span}\{\mathbf{w}_0, \mathbf{w}_{-1}, \ldots, \mathbf{w}_{k-1}, \mathbf{w}_{-k}\},\end{aligned}$$

thereby establishing equation (12). Next note that $\mathbf{w}_{k-1}$ is contained in

$$\begin{aligned}\operatorname{span}\{\mathbf{h}, U^{-1}\mathbf{h}, \ldots, U^{k-2}\mathbf{h}, U^{-k+1}\mathbf{h}, U^{k-1}\mathbf{h}\} &= \operatorname{span}\{\mathbf{w}_0, \mathbf{w}_{-1}, \ldots, \mathbf{w}_{k-2}, \mathbf{w}_{-k+1}, U^{k-1}\mathbf{h}\}\\ &= \operatorname{span}\{\tilde{\mathbf{v}}_{-1}, \tilde{\mathbf{v}}_0, \ldots, \tilde{\mathbf{v}}_{-k+1}, \tilde{\mathbf{v}}_{k-2}, U^{k-1}\mathbf{h}\},\end{aligned}$$

where the last equality is due to (12). A key observation is that when writing $\mathbf{w}_{k-1}$ down as a linear combination of the vectors $\tilde{\mathbf{v}}_{-1}, \tilde{\mathbf{v}}_0, \ldots, \tilde{\mathbf{v}}_{-k+1}, \tilde{\mathbf{v}}_{k-2}$ and $U^{k-1}\mathbf{h}$, the coefficient in the direction of $U^{k-1}\mathbf{h}$ is nonzero.

Hence, we obtain

$$\begin{aligned}\operatorname{span}\{\tilde{\mathbf{v}}_{-1}, \tilde{\mathbf{v}}_0, \ldots, \tilde{\mathbf{v}}_{-k}, \tilde{\mathbf{v}}_{k-1}\} &= \operatorname{span}\{\tilde{\mathbf{v}}_{-1}, \tilde{\mathbf{v}}_0, \ldots, \tilde{\mathbf{v}}_{-k}, U^{k-1}\mathbf{h}\}\\ &= \operatorname{span}\{\tilde{\mathbf{v}}_{-1}, \tilde{\mathbf{v}}_0, \ldots, \tilde{\mathbf{v}}_{-k}, \mathbf{w}_{k-1}\},\end{aligned}$$

where the second equality is justified by the observation made in the previous paragraph, inducing equality (13). Equality (14) can be deduced using the same reasoning as for equation (13) by interchanging the roles of $(\tilde{\mathbf{v}}_k)$ and $(\mathbf{w}_k)$.

□

Equalities (12), (13) and (14) are due to the crosses linking the sequences corresponding to $(\tilde{\mathbf{v}}_k)$ and $(\mathbf{w}_k)$, whereas equalities (15), (16) and (17) are due to the crosses linking the sequences corresponding to $(\mathbf{w}_k)$ and $(\mathbf{v}_k)$. The columns in Figure 8 determine the order in which the orthonormal vectors will be computed. Intuitively each cross in Figure 8 represents a coupled two term recurrence. Going through the pattern of crosses, the orthonormal vectors are computed columnwise. Note that $(\tilde{\mathbf{v}}_k)$ is only utilized to connect $(\mathbf{v}_k)$ and $(\mathbf{w}_k)$, and does not appear explicitly in the recurrences.

As an example for the application of Lemma 4, consider the computation of $\mathbf{v}_2$, $\tilde{\mathbf{v}}_1$ and $\mathbf{w}_{-2}$, i.e., the fourth column of orthonormal vectors to be computed in Figure 8. To compute $\mathbf{w}_{-2}$ one has to orthonormalize $U^{-2}\mathbf{h}$ against $\mathbf{w}_0$, $\mathbf{w}_{-1}$ and $\mathbf{w}_1$. Equation (14) implies that $U^{-2}\mathbf{h}$ can by replaced by $\tilde{\mathbf{v}}_{-2}$ without altering the resulting vector $\mathbf{w}_{-2}$, i.e., one can orthonormalize $\tilde{\mathbf{v}}_{-2}$ against $\mathbf{w}_0$, $\mathbf{w}_{-1}$ and $\mathbf{w}_1$ in order to obtain $\mathbf{w}_{-2}$. As $\tilde{\mathbf{v}}_{-2}$ is already orthogonal to

$$\operatorname{span}\{\mathbf{w}_0, \mathbf{w}_{-1}\} = \operatorname{span}\{\mathbf{h}, U^{-1}\mathbf{h}\} = \operatorname{span}\{\tilde{\mathbf{v}}_{-1}, \tilde{\mathbf{v}}_0\},$$

where the latter is due to (12), $\mathbf{w}_{-2}$ is the result of orthonormalizing $\tilde{\mathbf{v}}_{-2}$ against $\mathbf{w}_1$. Analogously, $\tilde{\mathbf{v}}_1$ which is part of the same cross structure as $\mathbf{w}_{-2}$ is the result of orthonormalizing $\mathbf{w}_1$ against $\tilde{\mathbf{v}}_{-2}$. The vector $\mathbf{v}_2$ is computed as $U\tilde{\mathbf{v}}_1$.

In general, by the same reasoning as in the previous paragraph, the crosses connecting the sequences corresponding to $(\tilde{\mathbf{v}}_k)$ and $(\mathbf{w}_k)$ give rise to the recurrence relations

$$\tilde{\mathbf{v}}_k\beta_k = \mathbf{w}_k + \tilde{\mathbf{v}}_{-k-1}\alpha_k,$$

$$\mathbf{w}_{-k-1}\beta_k = \tilde{\mathbf{v}}_{-k-1} + \mathbf{w}_k\bar{\alpha}_k,$$

or equivalently

$$\mathbf{v}_{k+1}\beta_k = U\mathbf{w}_k + \mathbf{v}_{-k}\alpha_k,$$

$$\mathbf{w}_{-k-1}\beta_k = U^{-1}\mathbf{v}_{-k} + \mathbf{w}_k\bar{\alpha}_k,$$

where

$$\alpha_k = -\langle U\mathbf{w}_k, \mathbf{v}_{-k}\rangle, \qquad \beta_k = \sqrt{1-|\alpha_k|^2}.$$

The crosses linking the sequences corresponding to $(\mathbf{v}_k)$ and $(\mathbf{w}_k)$ give rise to the recurrence relations

$$\mathbf{w}_{k+1}\epsilon_k = \mathbf{v}_{k+1} + \mathbf{w}_{-k-1}\delta_k,$$

$$\mathbf{v}_{-k-1}\epsilon_k = \mathbf{w}_{-k-1} + \mathbf{v}_{k+1}\bar{\delta}_k,$$

where

$$\delta_k = -\langle \mathbf{v}_{k+1}, \mathbf{w}_{-k-1}\rangle, \qquad \epsilon_k = \sqrt{1-|\delta_k|^2}.$$

Consequently, we have rediscovered the recurrence relations discussed in [15] for computing the bases $(\mathbf{v}_k)$ and $(\mathbf{w}_k)$ simultaneously.

Given the recurrence relations, one can see that the matrix of U with respect to the bases $(\mathbf{w}_k)$ and $(\mathbf{v}_k)$ is given by the matrix

$$X = \mathbf{V}^*U\mathbf{W} = \begin{bmatrix} -\alpha_0 & \beta_0 & & & & & \\ \beta_0 & \bar{\alpha}_0 & & & & & \\ & & -\alpha_1 & \beta_1 & & & \\ & & \beta_1 & \bar{\alpha}_1 & & & \\ & & & & -\alpha_2 & \beta_2 & \\ & & & & \beta_2 & \bar{\alpha}_2 & \\ & & & & & & \ddots \end{bmatrix},$$

and the matrix of the identity operator with respect to the bases $(\mathbf{v}_k)$ and $(\mathbf{w}_k)$ is given by the matrix

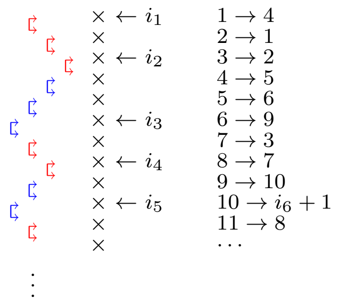


Fig. 9: Graphical representation of the swapping procedure.

$$
\begin{array}{lccccccccccc}
\text{sequence (18) :} & \mathbf{h} & U\mathbf{h} & U^2\mathbf{h} & U^{-1}\mathbf{h} & U^{-2}\mathbf{h} & U^{-3}\mathbf{h} & U^{3}\mathbf{h} & U^{4}\mathbf{h} & U^{-4}\mathbf{h} & U^{-5}\mathbf{h} & \ldots \\
\downarrow \times U^{-1} & & & & & & & & & & & \\
\text{1st auxiliary seq. :} & U^{-1}\mathbf{h} & \mathbf{h} & U\mathbf{h} & U^{-2}\mathbf{h} & U^{-3}\mathbf{h} & U^{-4}\mathbf{h} & U^{2}\mathbf{h} & U^{3}\mathbf{h} & U^{-5}\mathbf{h} & U^{-6}\mathbf{h} & \ldots \\
\text{2nd auxiliary seq. :} & U^{-1}\mathbf{h} & \mathbf{h} & U\mathbf{h} & U^{-2}\mathbf{h} & U^{-3}\mathbf{h} & U^{2}\mathbf{h} & U^{-4}\mathbf{h} & U^{3}\mathbf{h} & U^{-5}\mathbf{h} & U^{4}\mathbf{h} & \ldots \\
\text{3th auxiliary seq. :} & \mathbf{h} & U^{-1}\mathbf{h} & U\mathbf{h} & U^{-2}\mathbf{h} & U^{2}\mathbf{h} & U^{-3}\mathbf{h} & U^{3}\mathbf{h} & U^{-4}\mathbf{h} & U^{4}\mathbf{h} & U^{-5}\mathbf{h} & \ldots \\
\text{4th auxiliary seq. :} & \mathbf{h} & U\mathbf{h} & U^{-1}\mathbf{h} & U^{2}\mathbf{h} & U^{-2}\mathbf{h} & U^{-3}\mathbf{h} & U^{3}\mathbf{h} & U^{4}\mathbf{h} & U^{-4}\mathbf{h} & U^{-5}\mathbf{h} & \ldots \\
\text{sequence (18) :} & \mathbf{h} & U\mathbf{h} & U^{2}\mathbf{h} & U^{-1}\mathbf{h} & U^{-2}\mathbf{h} & U^{-3}\mathbf{h} & U^{3}\mathbf{h} & U^{4}\mathbf{h} & U^{-4}\mathbf{h} & U^{-5}\mathbf{h} & \ldots
\end{array}
$$

Fig. 10: The swapping procedure applied.

$$
Y = \mathbf{W}^* I \mathbf{V} = \begin{bmatrix} 1 & & & & & & & \\ & -\delta_0 & \varepsilon_0 & & & & & \\ & \varepsilon_0 & \bar{\delta}_0 & & & & & \\ & & & -\delta_1 & \varepsilon_1 & & & \\ & & & \varepsilon_1 & \bar{\delta}_1 & & & \\ & & & & & -\delta_2 & \varepsilon_2 & \\ & & & & & \varepsilon_2 & \bar{\delta}_2 & \\ & & & & & & & \ddots \end{bmatrix}.
$$

Hence, the matrix of U with respect to the bases $(\mathbf{v}_k)$ is given by XY, which is a matrix of pentadiagonal form, as already shown in Subsection 3.3.

4.2 The swapping procedure

In this section the technique of swapping elements as described in Subsection 4.1 will be generalized, with the intention of determing two term recurrences for arbitrary extended Krylov sequences for unitary matrices.

We will explain this by means of an example. Consider the pattern of rotations as depicted in Figure 9, which corresponds to the sequence

$$
\mathbf{h},\ U\mathbf{h},\ U^2\mathbf{h},\ U^{-1}\mathbf{h},\ U^{-2}\mathbf{h},\ U^{-3}\mathbf{h},\ U^{3}\mathbf{h},\ U^{4}\mathbf{h},\ U^{-4}\mathbf{h},\ U^{-5}\mathbf{h},\ U^{5}\mathbf{h},\ \ldots. \tag{18}
$$

The pattern of rotations as depicted in Figure 9 consists of four columns. From the sequence (18) four additional sequences will be constructed, denoted with auxiliary sequence 1 up to 4. These sequences are depicted in Figure 10. The first auxiliary sequence is obtained by multiplying sequence (18) with U^{-1}. Next we apply on the first auxiliary sequence what we will call the *swapping procedure*. The second auxiliary sequence is obtained by swapping neighbouring elements in the first auxiliary sequence whenever there is a rotation acting on the same positions in the first column of the pattern of rotations (9). The third auxiliary sequence is obtained by swapping neighbouring elements in the second auxiliary sequence whenever there is a rotation acting on the same positions in the second column of the pattern of rotations (9) and finally the fourth auxiliary sequence is obtained by swapping neighbouring elements in the third sequence whenever there is a rotation acting on the same positions in the third column of the pattern of rotations (9). The essence of the method is that by swapping neighbouring elements in the fourth (final) auxiliary sequence whenever there is a rotation acting on the same positions in the fourth (last) column of the pattern of rotations, one obtaines the original sequence (18). In order to prove this, consider Figure 9 where the result of the swapping procedure is shown, i.e., when applied to some sequence of vectors, the vector in the first position is transfered to the fourth position, the vector in the second position is transfered to the first position, and so on. Note that the pattern of rotations can be divided into two subsequences, a sequence corresponding to ascending exponents which is highlighted in red and a sequence corresponding to descending exponents which is highlighted in blue. As is depicted in Figure 9, if the ith rotation is highlighted in red, the ith element of the sequence is mapped to the jth element, where the jth rotation is the last rotation highlighted in red and preceding the ith one. If the ith rotation is highlighted in blue, the ith element of the sequence is mapped to the jth element, where the jth rotation is the first rotation highlighted in blue and following the ith one. In other words, applying the swapping procedure is indentical with multiplication by U. Therefore, applying the swapping procedure to the first auxiliary sequence, the latter being the original sequence (18) times U^{-1}, returns the original sequence (18).

Obviously, the same reasoning applies to any pattern of rotations and the associated number of sequences to be generated.

Remark 1 In order to compute an extended Krylov basis, the algorithm presented below introduces additional bases which are coupled with the latter basis and with each other by two term recurrences. These additional bases are created by the swapping procedure as described in this section. The number of additional bases is determined by the number of columns of the corresponding pattern of rotations. Sometimes the number of additional bases can be reduced by rewriting the pattern of rotations. Consider Figure 11. On the left a rotation pattern is shown consisting of four columns, whereas on the right a pattern of rotations is depicted consisting of only three columns, though both patterns represent the same matrix. Therefore, applying the algorithm described below on the left pattern will give rise to the computation of four additional bases and a factorization of the corresponding extended Hessenberg matrix comprised of four factor matrices, whereas applying the same algorithm on the left pattern will give rise to the computation of three additional bases and a factorization of the same extended Hessenberg matrix comprised of three factor matrices.

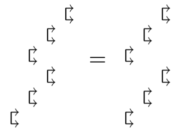


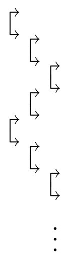
Fig. 11: Two patterns of rotations representing the same unitary matrix.

4.3 Two periodic pattern of rotations

Our aim is to apply the swapping procedure of Subsection 4.2 to arbitrary extended Krylov sequences. As an example to describe our method, we orthonormalize the sequence

$$\mathbf{h}, U\mathbf{h}, U^2\mathbf{h}, U^{-1}\mathbf{h},\ U^{-2}\mathbf{h},\ U^3\mathbf{h},\ U^4\mathbf{h},\ U^{-3}\mathbf{h},\ U^{-4}\mathbf{h},\ \ldots,$$

of which the position vector corresponds to the pattern of rotations (19).



(19)

The pattern (19) consists of three columns, hence we will simultaneously orthonormalize four sequences giving rise to the bases $(\mathbf{v}_k)$, $(\tilde{\mathbf{v}}_k)$, $(\mathbf{w}_k)$ and $(\mathbf{u}_k)$. First the sequence corresponding to $(\tilde{\mathbf{v}}_k)$ is constructed by multiplying the vectors of the sequence corresponding to $(\mathbf{v}_k)$ with U^{-1}. The sequence corresponding to $(\mathbf{w}_k)$ can be derived from that of $(\tilde{\mathbf{v}}_k)$ by swapping neighbouring elements whenever there is a rotation acting on the same positions in the first column. The sequence corresponding to $(\mathbf{u}_k)$ can be derived from that of $(\mathbf{w}_k)$ be swapping neighbouring elements whenever there is a rotation acting on the same positions in the second column. And finally, the sequence corresponding to $(\mathbf{v}_k)$ can be derived from that of $(\mathbf{u}_k)$ be swapping neighbouring elements whenever there is a rotation acting on the same positions in the third column. This links all three sequences in a cyclic way and is depicted in Figure 12. Note that the arrows in Figure 12 follow the same pattern as the rotations in (19) but mirrored.

$$\begin{array}{llllllllll}
\mathbf{v}: & \mathbf{h} & U\mathbf{h} & U^2\mathbf{h} & U^{-1}\mathbf{h} & U^{-2}\mathbf{h} & U^3\mathbf{h} & U^4\mathbf{h} & U^{-3}\mathbf{h} & U^{-4}\mathbf{h} & \ldots \\
\downarrow \times U^{-1} & & & & & & & & & & \\
\tilde{\mathbf{v}}: & U^{-1}\mathbf{h} & \mathbf{h} & U\mathbf{h} & U^{-2}\mathbf{h} & U^{-3}\mathbf{h} & U^2\mathbf{h} & U^3\mathbf{h} & U^{-4}\mathbf{h} & U^{-5}\mathbf{h} & \ldots \\
\mathbf{w}: & \mathbf{h} & U^{-1}\mathbf{h} & U\mathbf{h} & U^{-2}\mathbf{h} & U^2\mathbf{h} & U^{-3}\mathbf{h} & U^3\mathbf{h} & U^{-4}\mathbf{h} & U^4\mathbf{h} & \ldots \\
\mathbf{u}: & \mathbf{h} & U\mathbf{h} & U^{-1}\mathbf{h} & U^2\mathbf{h} & U^{-2}\mathbf{h} & U^3\mathbf{h} & U^{-3}\mathbf{h} & U^4\mathbf{h} & U^{-4}\mathbf{h} & \ldots \\
\mathbf{v}: & \mathbf{h} & U\mathbf{h} & U^2\mathbf{h} & U^{-1}\mathbf{h} & U^{-2}\mathbf{h} & U^3\mathbf{h} & U^4\mathbf{h} & U^{-3}\mathbf{h} & U^{-4}\mathbf{h} & \ldots
\end{array}$$

Fig. 12: Links between the sequences corresponding to the bases (v_k), (w_k) and (u_k).

Using Figure 12, similar equalities between the spans of several subspaces can be deduced as in Lemma 4. Again the crosses connecting the different vectors contain all necessary information. Suppose we have an orthonormal vector $\mathbf{y}$ of which the underlying product of a power of U with $\mathbf{h}$ is located in the right column of a cross. Then the span of all vectors in the corresponding basis computed before $\mathbf{y}$, including $\mathbf{y}$ itself, remains the same if $\mathbf{y}$ is replaced by the orthonormal vector which is connected to $\mathbf{y}$ by the cross. We do not write down these equalities since it would make our discussion unnecessarily complicated (due to the heavy use of indices) and would not contribute to the insight Figure 12 already provides us. Instead, we rely on Figure 13, which gives a schematic view on the order of computation of the orthonormal vectors.

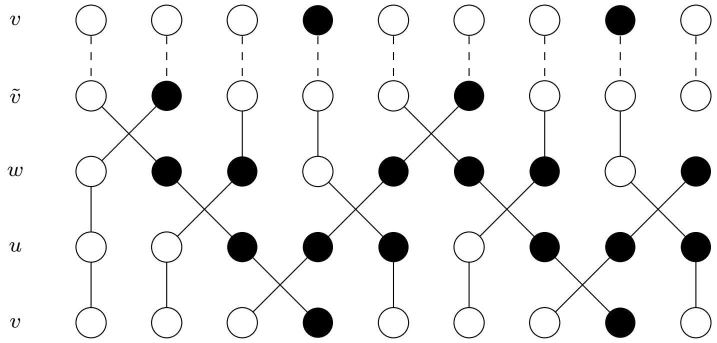


Fig. 13: Schematic view on the order of computation of the orthonormal vectors.

Figure 13 should be read together with Figure 12. The orthonormal vectors are computed column by column. In each column vectors which are located in a black disk are orthonormalized first. Vectors in the same column which are connected to each other with a solid line result in the same orthonormal vector. The dashed line represents the relation $\tilde{\mathbf{v}}_k = U^{-1}\mathbf{v}_{k+1}$.

To illustrate this, suppose we want to compute the fifth column in Figure 12, i.e., the orthonormal vectors $\mathbf{v}_{-2}$, $\tilde{\mathbf{v}}_{-3}$, $\mathbf{w}_2$ and $\mathbf{u}_{-2}$. The vectors $\mathbf{w}_2$ and $\mathbf{u}_{-2}$ are marked with a black circle in Figure 13 and are calculated first. The vector $\mathbf{w}_2$ is the result of orthonormalizing $U^2\mathbf{h}$ against $\mathbf{w}_0$, $\mathbf{w}_{-1}$, $\mathbf{w}_1$ and $\mathbf{w}_{-2}$. As $\mathbf{u}_2$ is a linear combination of $\mathbf{h}$, $U\mathbf{h}$, $U^{-1}\mathbf{h}$ and $U^2\mathbf{h}$ with a nonzero component in the direction of $U^2\mathbf{h}$, $U^2\mathbf{h}$ can be replaced by $\mathbf{u}_2$ without altering the resulting vector $\mathbf{w}_2$, i.e., the vector $\mathbf{w}_2$ is the result of orthonormalizing $\mathbf{u}_2$ against $\mathbf{w}_0$, $\mathbf{w}_{-1}$, $\mathbf{w}_1$ and $\mathbf{w}_{-2}$. Since $\mathbf{u}_2$ is already orthogonal to

$$\operatorname{span}\{\mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_{-1}\} = \operatorname{span}\{\mathbf{h}, U\mathbf{h}, U^{-1}\mathbf{h}\} = \operatorname{span}\{\mathbf{w}_0, \mathbf{w}_{-1}, \mathbf{w}_1\},$$

$\mathbf{w}_2$ is the result of orthonormalizing $\mathbf{u}_2$ against $\mathbf{w}_{-2}$. Similarly, $\mathbf{u}_{-2}$ is the result of orthonormalizing $\mathbf{w}_{-2}$ against $\mathbf{u}_2$. One obtains the recurrence relations

$$\mathbf{w}_2\varepsilon_1 = \mathbf{u}_2 + \mathbf{w}_{-2}\delta_1,$$

$$\mathbf{u}_{-2}\varepsilon_1 = \mathbf{w}_{-2} + \mathbf{u}_2\bar{\delta}_1,$$

where

$$\delta_1 = -\langle \mathbf{u}_2, \mathbf{w}_{-2}\rangle, \qquad \varepsilon_1 = \sqrt{1 - |\delta_1|^2}.$$

The vector $\mathbf{v}_{-2}$ is marked with a white circle and connected by a solid line with $\mathbf{u}_{-2}$. At the same time it is connected by a dashed line with $\tilde{\mathbf{v}}_{-3}$. The vector $\mathbf{v}_{-2}$ is the result of orthonormalizing $U^{-2}\mathbf{h}$ against

$$\operatorname{span}\{\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_{-1}\} = \operatorname{span}\{\mathbf{h}, U\mathbf{h}, U^2\mathbf{h}, U^{-1}\mathbf{h}\},$$

whereas the vector $\mathbf{u}_{-2}$ is the result of orthonormalizing $U^{-2}\mathbf{h}$ against

$$\operatorname{span}\{\mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_{-1}, \mathbf{u}_2\} = \operatorname{span}\{\mathbf{h}, U\mathbf{h}, U^{-1}\mathbf{h}, U^2\mathbf{h}\}.$$

The latter implies that $\mathbf{v}_{-2}$ and $\mathbf{u}_{-2}$ are obtained by orthonormalizing $U^{-2}\mathbf{h}$ against the same subspace, implying that both vectors are equal up to a unimodular factor, which we choose to be one. Therefore, one obtains

$$\mathbf{v}_{-2} = \mathbf{u}_{-2}.$$

Finally, by construction one has

$$\tilde{\mathbf{v}}_{-3} = U^{-1}\mathbf{v}_{-2}.$$

In this way one proceeds columnwise, computing the bases $(\mathbf{v}_k)$, $(\mathbf{w}_k)$ and $(\mathbf{u}_k)$ simultaneously.

The matrix of U with respect to the bases $(\mathbf{w}_k)$ and $(\mathbf{v}_k)$ is given by the matrix

$$X = \mathbf{V}^* U \mathbf{W} = \begin{bmatrix} -\alpha_0 & \beta_0 & & & & & & \\ \beta_0 & \bar{\alpha_0} & & & & & & \\ & & 1 & & & & & \\ & & & 1 & & & & \\ & & & & -\alpha_1 & \beta_1 & & \\ & & & & \beta_1 & \bar{\alpha}_1 & & \\ & & & & & & 1 & \\ & & & & & & & \ddots \end{bmatrix},$$

the matrix of the identity operator with respect to the bases $(\mathbf{u}_k)$ and $(\mathbf{w}_k)$ is given by the matrix

$$Y = \mathbf{W}^* I \mathbf{U} = \begin{bmatrix} 1 & & & & & & & \\ & -\delta_0 & \varepsilon_0 & & & & & \\ & \varepsilon_0 & \bar{\delta_0} & & & & & \\ & & & -\delta_1 & \varepsilon_1 & & & \\ & & & \varepsilon_1 & \bar{\delta}_1 & & & \\ & & & & & -\delta_2 & \varepsilon_2 & \\ & & & & & \varepsilon_2 & \bar{\delta}_2 & \\ & & & & & & & \ddots \end{bmatrix},$$

and the matrix of the identity operator with respect to the bases $(\mathbf{v}_k)$ and $(\mathbf{u}_k)$ is given by the matrix

$$Z = \mathbf{U}^* I \mathbf{V} = \begin{bmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & -\xi_0 & \gamma_0 & & & \\ & & \gamma_0 & \bar{\xi}_0 & & & \\ & & & & 1 & & \\ & & & & & 1 & \\ & & & & & & \ddots \end{bmatrix}.$$

The product XYZ is a matrix of extended Hessenberg form corresponding to the position vector $\mathbf{p} = [\ell, \ell, r, r, \ell, \ell, \ldots]$.

4.4 Arbitrary rotation patterns and algorithm

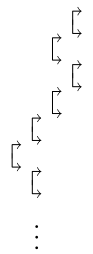
In this section a new algorithm is presented for computing an orthonormal extended Krylov basis.

Algorithm 1 returns an orthonormal basis $(\mathbf{v}_k)$ for an arbitrary extended Krylov subspace, given a unitary matrix U, a starting vector $\mathbf{h}$ and a position vector $\mathbf{p}$ determining the ordering of the extended Krylov sequence. Together with the basis vectors $(\mathbf{v}_k)$ the corresponding extended Hessenberg matrix H, i.e., the projection of U on the extended Krylov subspace, is returned. When examining Algorithm 1, note that to compute the upper left $m \times m$ submatrix of H correctly, a position vector of length m is required.

We will now explain the mechanics of the algorithm by means of an example. Consider the extended Krylov sequence

$$\mathbf{h},\ U^{-1}\mathbf{h},\ U\mathbf{h},\ U^{-2}\mathbf{h},\ U^{-3}\mathbf{h},\ U^{-4}\mathbf{h},\ U^{2}\mathbf{h},\ \ldots,$$

corresponding to the position vector $\mathbf{p} = [r, \ell, r, r, r, \ell, \ldots]$ and therefore to the pattern of rotations (20).



(20)

The pattern (20) consists of four columns, hence in addition to the sequences $(\mathbf{w}_k)$ and $(\mathbf{u}_k)$ an extra sequence $(\mathbf{z}_k)$ is needed. This is visualized in Figure 14.

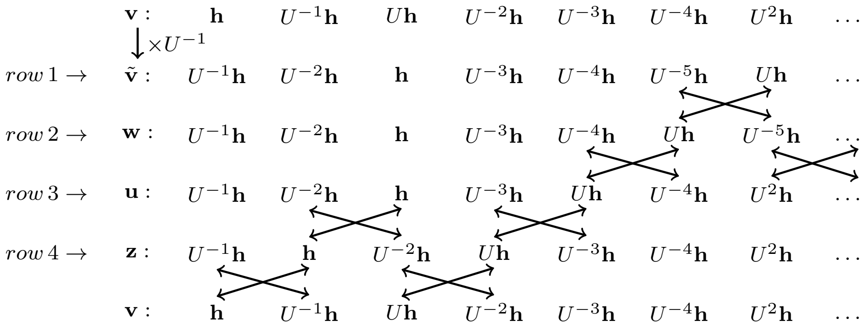
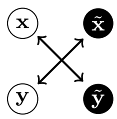


Fig. 14: Links between the sequences corresponding to $(\mathbf{v}_k)$, $(\mathbf{w}_k)$, $(\mathbf{u}_k)$ and $(\mathbf{z}_k)$.

Figure 15 shows the order in which the orthonormal vectors are computed. Given this and the examples of the previous subsections, the building blocks for a general algorithm can be constructed.

Vectors which are part of a cross structure in Figures 14 and 15 are computed by the recurrence relations



$$\tilde{\mathbf{x}}\beta = \mathbf{y} + \mathbf{x}\alpha,$$

$$\tilde{\mathbf{y}}\beta = \mathbf{x} + \mathbf{y}\bar{\alpha},$$

where $\alpha = -\langle \mathbf{y}, \mathbf{x} \rangle$ and $\beta = \sqrt{1 - |\alpha|^2}$. The coefficients α and β are stored as they are part of the sparse factorization of the corresponding extended Hessenberg matrix.

Orthonormal vectors which are connected with a solid line in Figure 15 are the same. The dashed line indicates multiplication with U, i.e., for each $k \in \mathbb{Z}$, $\mathbf{v}_k = U\tilde{\mathbf{v}}_{k-1}$. As the vectors indicated with a black disk are computed first, the vectors indicated with a white disk are found at once. This allows us to run through Figure 15 column by column, computing the bases $(\mathbf{v}_k)$, $(\mathbf{w}_k)$, $(\mathbf{u}_k)$ and $(\mathbf{z}_k)$ simultaneously. Note that as our main interest is to compute the basis $(\mathbf{v}_k)$, we do not need to store the whole bases $(\tilde{\mathbf{v}}_k)$, $(\mathbf{w}_k)$, $(\mathbf{u}_k)$ and $(\mathbf{z}_k)$, but only the most recently computed vectors of the latter bases as these are the vectors necessary to compute the next basis vector of $(\mathbf{v}_k)$.

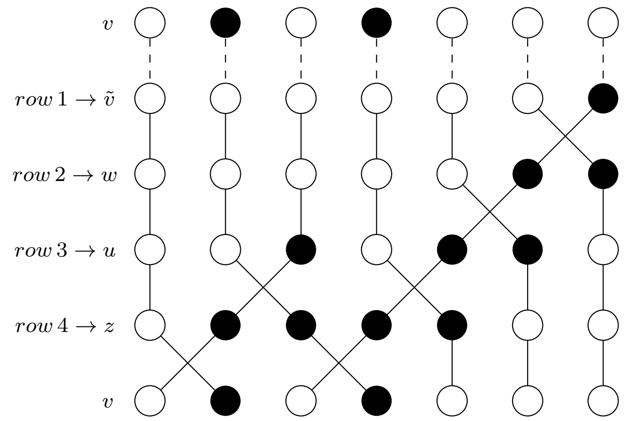


Fig. 15: Schematic view on the order of computation of the orthonormal vectors.

We will now clarify some of the notation used in Algorithm 1. In contrast to the rest of the manuscript, the orthonormal basis vectors $(\mathbf{v}_k)$ are indexed linearly, i.e., k denotes the order in which the vectors are computed. In each iteration only the most recently computed vectors of the auxiliary bases need to be stored, these are denoted by the vectors $\mathbf{w}_1, \ldots, \mathbf{w}_{width}$. In the example above $width = 4$ and $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$, $\mathbf{w}_4$ correspond to the most recently computed vector of the bases $(\tilde{\mathbf{v}}_k)$, $(\mathbf{w}_k)$, $(\mathbf{u}_k)$ and $(\mathbf{z}_k)$ respectively. Note that the variable $width$ corresponds with the width of the pattern (20). Finally, the vector row is used to keep track of the current location in the pattern. More specifically, $row(i)$ indicates the row number of the upper left element of the ith cross. In the example of Figure 15 we have $row = [4, 3, 4, 3, 2, 1, \ldots]$. The vector row is easily obtained from the position vector $\mathbf{p}$.

Algorithm 1 clearly exhibits a computational advantage with respect to the classical Arnoldi algorithm, as in each iteration only one inner product needs to be computed. In order to gain this computational advantage, w extra bases need to be computed simultaneously with the original extended Krylov basis, where w is the width of the corresponding pattern of rotations. The width w of this pattern is at most equal to the number of orthonormal vectors to be retrieved, i.e., when the position vector $\mathbf{p}$ consists entirely of ℓ's or entirely of r's. As we only need to store the most recently computed orthonormal vectors of this extra bases, the memory cost of Algorithm 1 is at worst twice as high as the memory cost of Arnoldi. Another advantage of Algorithm 1 is the computation of the associated extended Hessenberg matrix H, which is returned in factored form, the factors beiing sparse block matrices consisting of an alternating sequence of 2×2 unitary rotations and identity blocks.

In case the position vector $\mathbf{p}$ consists of an entire sequence of ℓ's, the algorithm will compute an orthonormal Krylov basis for a unitary matrix. For this case, the reader might be interested in the isometric Arnoldi algorithm [6, 7, 16], which is also based on coupled two-term recurrences.

5 Longer recurrences

In this section another type of recurrence is constructed where we simultaneously compute one additional basis $(\mathbf{w}_k)$ together with the orginal basis $(\mathbf{v}_k)$. Doing this, one obtains longer recurrences, of which the length depends on the differences $i_{j+1} - i_j$. As an example, we consider the sequence

$$\mathbf{h},\ U\mathbf{h},\ U^2\mathbf{h},\ U^3\mathbf{h},\ U^{-1}\mathbf{h},\ U^{-2}\mathbf{h},\ U^4\mathbf{h},\ U^5\mathbf{h},\ \ldots,$$

corresponding to the pattern (21).

Algorithm 1 Algorithm for computing orthonormal extended Krylov bases.

Require: unitary matrix U, starting vector $\mathbf{h}$, position vector $\mathbf{p}$

▷ Initialization

$\mathbf{v}_1 \leftarrow \mathbf{h}/||\mathbf{h}||_2$;
for $i = 1 \rightarrow row(1)$ **do**
  $\mathbf{w}_i \leftarrow U^*\mathbf{v}_1$;
end for
for $i = row(1)+1 \rightarrow width$ **do**
  $\mathbf{w}_i \leftarrow \mathbf{v}_1$;
end for
for $i = 1 \rightarrow width$ **do**
  $H_i = I_{length\,\mathbf{p}+1}$
end for

▷ Update vectors located in black disks

for $k = 2 \rightarrow length\,\mathbf{p}+1$ **do**
  if $row(k-1) = width$ **then**
    $\mathbf{x} \leftarrow \mathbf{w}_{row(k-1)}$;
    $\mathbf{y} \leftarrow \mathbf{v}_{k-1}$;
    $\alpha \leftarrow -\overline{\langle \mathbf{y}, \mathbf{x} \rangle}$;
    $\beta \leftarrow \sqrt{1-\alpha\bar{\alpha}}$;
    $\mathbf{w}_{row(k-1)} \leftarrow (y+\alpha x)/\beta$;
    $\mathbf{v}_k \leftarrow (x+\bar{\alpha}y)/\beta$;
  else
    $\mathbf{x} \leftarrow \mathbf{w}_{row(k-1)}$;
    $\mathbf{y} \leftarrow \mathbf{w}_{row(k-1)+1}$;
    $\alpha \leftarrow -\overline{\langle \mathbf{y}, \mathbf{x} \rangle}$;
    $\beta \leftarrow \sqrt{1-\alpha\bar{\alpha}}$;
    $\mathbf{w}_{row(k-1)} \leftarrow (\mathbf{y}+\alpha\mathbf{x})/\beta$;
    $\mathbf{w}_{row(k-1)+1} \leftarrow (\mathbf{x}+\bar{\alpha}\mathbf{y})/\beta$;
  end if

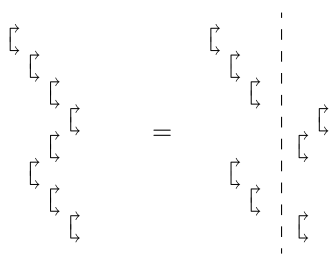
▷ Update vectors located in white disks

  if $row(k) > row(k-1)$ **then**
    for $i = 1 \rightarrow row(k-1)-1$ **do**
      $\mathbf{w}_i \leftarrow \mathbf{w}_{row(k-1)}$;
    end for
    $\mathbf{v}_k \leftarrow U\mathbf{w}_1$;
    for $i = row(k-1)+2 \rightarrow width$ **do**
      $\mathbf{w}_i = \mathbf{v}_k$;
    end for
  else
    if $row(k-1) \neq width$ **then**
      for $i = row(k-1)+2 \rightarrow width$ **do**
        $\mathbf{w}_i \leftarrow \mathbf{w}_{row(k-1)+1}$;
      end for
      $\mathbf{v}_k \leftarrow \mathbf{w}_{width}$;
      for $i = 1 \rightarrow row(k-1)-1$ **do**
        $\mathbf{w}_i \leftarrow U^* v_k$;
      end for
    end if
  end if

▷ Update Hessenberg factorization

$$\begin{pmatrix} \left(H_{row(k-1)}\right)_{k-1,k-1} & \left(H_{row(k-1)}\right)_{k-1,k} \\ \left(H_{row(k-1)}\right)_{k,k-1} & \left(H_{row(k-1)}\right)_{k,k} \end{pmatrix} \leftarrow \begin{pmatrix} -\alpha & \beta \\ \beta & \bar{\alpha} \end{pmatrix}$$

end for
$H \leftarrow H_1 H_2 \ldots H_{width}$
return $(\mathbf{v}_k)$, H



(21)

The pattern (21) is divided into two parts, the left part consisting of descending sequences, the right part consisting of ascending sequences. Figure 16 depicts the construction of the auxiliary basis $(\mathbf{w}_k)$.

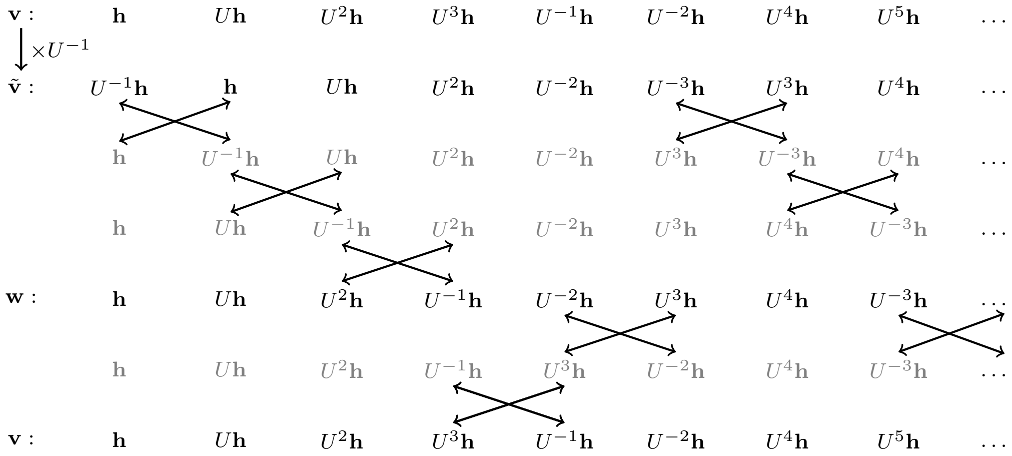


Fig. 16: Links between the sequences corresponding to the bases $(\mathbf{v}_k)$ and $(\mathbf{w}_k)$.

It is easily seen that the matrix of U with respect to the bases $(\mathbf{w}_k)$ and $(\mathbf{v}_k)$ is of the form

$$X = \mathbf{V}^* U \mathbf{W} = \qquad , \tag{22}$$

whereas the matrix of the identity operator with respect to the bases $(\mathbf{w}_k)$ and $(\mathbf{v}_k)$ is of the form

$$Y^* = \mathbf{V}^* \mathbf{W} = \qquad . \tag{23}$$

Note how the structure of the matrices X and Y is captured in the pattern (21). The product XY is the extended Hessenberg matrix corresponding to the position vector $\mathbf{p} = [\ell, \ell, \ell, r, r, \ell, \ell, \ldots]$.

For the ease of notation we will denote by $\mathbf{v}_i$, $\mathbf{w}_i$ the ith computed basis vector of $(\mathbf{v}_k)$, $(\mathbf{w}_k)$ respectively. It is also assumed that the position vector $\mathbf{p}$ starts with the symbol ℓ, which consistent with the above example. The following recurrences are retrieved. For $i_j \le i < i_{j+1}$, j odd, we obtain

$$\beta \mathbf{v}_{i+1} = U\mathbf{w}_i - \sum_{k=i_j}^{i} \alpha_k \mathbf{v}_k, \qquad \alpha_k = \langle U\mathbf{w}_i, \mathbf{v}_k \rangle,\ \beta = \sqrt{1 - \sum_{k=i_j}^{i} |\alpha_k|^2}$$

and

$$\mathbf{w}_{i+1} = \mathbf{v}_{i+1}, \qquad i_j < i < i_{j+1} - 1.$$

The vector $\mathbf{w}_{i_{j+1}}$ is computed as

$$\beta \mathbf{w}_{i_{j+1}} = U^* \mathbf{v}_{i_j} - \sum_{k=i_j}^{i_{j+1}-1} \alpha_k \mathbf{w}_k, \qquad \alpha_k = \langle U^* \mathbf{v}_{i_j}, \mathbf{w}_k \rangle,\ \beta = \sqrt{1 - \sum_{k=i_j}^{i_{j+1}-1} |\alpha_k|^2}.$$

For $i_j \leq i < i_{j+1}$, j even, we obtain

$$\beta \mathbf{v}_{i+1} = \mathbf{w}_i - \sum_{k=i_j}^{i} \alpha_k \mathbf{v}_k, \qquad \alpha_k = \langle \mathbf{w}_i, \mathbf{v}_k \rangle,\ \beta = \sqrt{1 - \sum_{k=i_j}^{i} |\alpha_k|^2}$$

and

$$\mathbf{w}_{i+1} = \mathbf{v}_{i+1}, \qquad i_j < i < i_{j+1} - 1.$$

The vector $\mathbf{w}_{i_{k+1}}$ is computed as

$$\beta \mathbf{w}_{i_{j+1}} = \mathbf{v}_{i_j} - \sum_{k=i_j}^{i_{j+1}-1} \alpha_k \mathbf{w}_k, \qquad \alpha_k = \langle \mathbf{v}_{i_j}, \mathbf{w}_k \rangle,\ \beta = \sqrt{1 - \sum_{k=i_j}^{i_{j+1}-1} |\alpha_k|^2}.$$

Note that the Hessenberg blocks in X and Y^* are unitary and therefore have a semiseparable structure. One can exploit this fact to shorten the above recurrences, e.g., [3].

6 Conclusion

We established the general structure of the projection of a matrix onto an extended Krylov subspace as a matrix containing a sequence of diagonal blocks which are alternatingly of Hessenberg and inverse Hessenberg form. This allowed us to derive the recurrences for Hermitian matrices as established in [8], without the use of orthonormal polynomial relations, nor was it necessary to discern between the positive and negative definite case. Furthermore, we investigated the structure of unitary extended Hessenberg matrices and derived corresponding recurrences, including the CMV-form and its associated coupled two-term recurrence as was established in [15].

References

1. Athanasios C. Antoulas. *Approximation of Large-Scale Dynamical Systems.* Society for Industrial and Applied Mathematics, 2005.
2. Ruyman Cruz Barroso and S. Delvaux. Orthogonal Laurent polynomials on the unit circle and snake-shaped matrix factorizations. *Journal of Approximation Theory*, 161(1):65–87, November 2009.
3. T. Barth and T. Manteuffel. Multiple Recursion Conjugate Gradient Algorithms Part I: Sufficient Conditions. *SIAM Journal on Matrix Analysis and Applications*, 21(3):768–796, 2000.
4. P. Benner, V. Mehrmann, and D.C. Sorensen. *Dimension Reduction of Large-Scale Systems*, volume 45. Lecture Notes in Computational Science and Engineering, 2005.
5. M. J. Cantero, L. Moral, and L. Velazquez. Five-diagonal matrices and zeros of orthogonal polynomials on the unit circle. *Linear Algebra and its Applications*, 362:29–56, March 2003.
6. W.B. Gragg. The QR-algorithm for unitary Hessenberg matrices. *Journal of Computational and Applied Mathematics*, 16:1–8, 1986.
7. W.B. Gragg. Positive definite Toeplitz matrices, the Arnoldi process for isometric operators, and Gaussian quadrature on the unit circle. *Journal of Computational and Applied Mathematics*, 46(1-2):183–198, 1993.
8. C. Jagels and L. Reichel. The extended Krylov subspace method and orthogonal Laurent polynomials. *Linear Algebra and its Applications*, 431(3):441–458, 2009.
9. L. Knizhnerman and V. Simoncini. A new investigation of the extended Krylov subspace method for matrix function evaluations. *Numerical Linear Algebra with Applications*, 17(4):615–638, 2010.
10. L. Knizhnerman and V. Simoncini. Convergence analysis of the extended Krylov subspace method for the Lyapunov equation. *Numerische Mathematik*, 118(3):567–586, 2011.
11. B. Simon. CMV matrices: Five years after. *Journal of Computational and Applied Mathematics*, 208(1):120–154, 2007.

12. R. Vandebril. Chasing bulges or rotations? A metamorphosis of the QR-algorithm. *SIAM Journal on Matrix Analysis and Applications*, 32:217–247, 2011.
13. R. Vandebril, M. Van Barel, and N. Mastronardi. *Matrix Computations and Semiseparable Matrices, Volume I: Linear Systems.* Johns Hopkins University Press, Baltimore, Maryland, USA, 2008.
14. R. Vandebril and D. S. Watkins. A generalization of the multishift QR algorithm. *SIAM Journal on Matrix Analysis and Applications*, 33(3):759–779, 2012.
15. D. S. Watkins. Some perspectives on the eigenvalue problem. *SIAM Review*, 35(3):430–471, 1993.
16. D.S. Watkins. Unitary orthogonalization processes. *Journal of Computational and Applied Mathematics*, 86(1):335–345, 1997.